Historic, archived document

Do not assume content reflects current scientific knowledge, policies, or practices.

cserve
280.39
34Am

AGRICULTURAL ECONOMICS RESEARCH

A Journal of Economic and Statistical Research in the
United States Department of Agriculture and Cooperating Agencies

U. S. DEPART... AGRICULTURE

Volume XII JULY 1960 Number 3

Observation and Audit Techniques for Measuring Retail Sales

By Earl E. Houseman and Benjamin Lipstein

Over the last few years several groups have experimented with the use of "observation" techniques for measuring the retail movement of fresh produce during a short period of time—a few days or a week. Some investigators have stated that the observation technique offers unique statistical as well as collection advantages. The present investigation was undertaken to compare the relative efficiency of the observation technique with the standard retail audit method, and to develop an efficient system of using observation techniques for measuring retail sales. This study, involving only apples and pears, was conducted in Philadelphia by Audits and Surveys Company, Inc., under contract with the Agricultural Marketing Service.

THE STANDARD retail auditing method consists of (1) taking an opening inventory of the product at the beginning of the sales period; (2) recording at the end of the period, which may be a week, a month or 2 months, the retailer's purchases of the product for the period audited; (3) taking a closing inventory; and (4) computing sales by means of the standard formula:

Opening Inventory + Purchases − Closing Inventory = Sales

The accuracy of the retail audit method of collecting retail sales data is a function of the accuracy of counting inventories, the care with which the retailer keeps purchase records, and records of returns and spoilage. Loss or omission of purchase records causes an understatement of sales; omission of returns and spoilage results in an excess of sales. At times, a failure in retailer records or an inaccurate inventory can be recognized when the standard equation indicates a negative sale. For example, if the closing inventory is greater than the opening inventory and no purchases are recorded, the record of indicated sales is negative. This is more likely to occur with a nonperishable product that has a high ratio of inventory to sales than with a perishable product such as apples or pears, because purchases are frequent and inventories are only a small fraction of total sales. Aside from the reasonableness of the sales data, there is no unique means for measuring respondent failure.

In addition to sales, the retail audit method provides data on inventories and purchases, information that cannot be derived by the observation method—information that may be important in the analysis of the market.

Observation Method

In the observation method, an interviewer records the quantity and unit price for each sale made during specified hours in specified stores. He can report whether a transaction was a bulk or prepackaged sale and state the physical condition of the merchandise. "Exposure" of the products can be determined by a count of customers

Reprinted by the Agricultural Marketing Service from AGRICULTURAL ECONOMICS RESEARCH, Vol. XII No. 3 July 1960

AMS-403

passing the counter as well as the number of actual transactions. This provides important and useful information on the number of customers and the size of individual purchases. Note that such factors as condition of the product and number of customers "exposed" are not available from either audit or consumer panels.

The audit method permits ascertaining the price of commodities only during the time of the audit, generally a slow sales period. On the other hand, if observation hours are selected at random, an average price can be computed that is an estimate of the average, either weighted by quantity or unweighted, for all transactions during the week.

Description of the Survey

There were two phases to the study, involving different marketing seasons. The plan followed in the second phase—that is, the second marketing season, October 1956 through April 1957—was the result of findings from the first phase, that is, the work conducted during the first marketing season, October 1955 through April 1956.

A preliminary survey was conducted just prior to the first marketing season. Its purpose was (1) to identify food stores that had fresh produce counters and two or more checkout counters and (2) to obtain information on their annual dollar volume, and when slow, medium, and busy sales periods occur during a week. In this preliminary study, about 650 stores were enumerated. A combined area and list type of sample was used. The sample of observation and audit stores described below was selected from the stores in the preliminary survey having self-service, fresh produce counters, and two or more checkout counters.

Throughout the first marketing season, both audit and observation measurements were taken in each sample store. Each store was audited once a week using standard auditing procedures on Monday, Tuesday, or Wednesday, to provide weekly estimates of sales. During each of the selected observation hours, one hour for each of the sample stores, an enumerator was stationed at the fresh produce counter to record the desired information about all customer purchases of apples and pears, including a count of customers passing the fresh produce counter. These observations were limited to Thursday, Friday, and Saturday.

Two equivalent stratified random samples of 30 stores (samples A and B) were selected from the list of stores enumerated in the preliminary survey. Stores on the list were stratified by dollar volume classes, and the sample was allocated to strata in proportion to number of stores. The sample of hours for observation was a random sample, with stratification by slow, medium, and busy periods and a restriction that only one hour for observation be drawn from each store. Two samples of 30 stores were used rather than a single sample of 60 stores, as one purpose of the study was to provide some information on precision, in a form understandable to persons with no training in statistical methods, by comparing results from the two samples.

Variance of Change

It was evident that the variance of observation estimates would be greater than the variance of audit estimates because the observation technique involves the subsampling of time within stores. That is, the sampling error for the observation estimates was comprised of variation between stores and variation between observation periods within stores, whereas the audit estimates were subject only to variation between stores. However, to make a choice between the two methods, information was needed on the magnitude of the differences in sampling error as well as on costs and other factors. Another question was, What is the optimum length of observation period? Consequently, the observers were instructed to keep separate records for the first and second half-hour of each hour's observation so the efficiency of half-hour and one-hour observation periods could be compared. The question of length of observation period is discussed later in more detail.

The two methods of measuring week-to-week change were compared using ratio estimates of the form $\frac{\Sigma X}{\Sigma Y}$, in which X represents audit or observation data on sales for the current week and Y represents the corresponding sales data for the previous week. For one pair of weeks for each of 4 months, the variance of the ratio was computed for the audit and observation methods. Square roots of the averages of these four variances are presented in table 1. These results are for a sample of 30 stores and a time schedule for taking observations that remains unchanged from week to week.

TABLE 1.—*Standard errors of week-to-week ratios for a sample of 30 stores*[1]

Commodity	Standard error		
	Observation		Audit
	1 hour	½ hour	
All apples	0. 12	0. 15	0. 09
All Eastern apples	. 13	--------	. 09
Eastern Delicious	. 22	--------	. 11
Eastern McIntosh	. 32	--------	. 17
Eastern Stayman	. 33	--------	. 15
All Western apples	. 24	--------	. 20
Western Delicious	. 31	--------	. 23
Pears	. 21	. 25	. 11

[1] This table is a reproduction of table 8 in "Measuring Weekly Changes in the Movement of Apples and Pears, 1955–56 Marketing Season" by Houseman, Hoofnagle, and Mullen—a preliminary report issued in February 1957 by the Agricultural Marketing Service.

From table 1, an important point to be noted is that the difference in efficiency is least for all apples. In general, as the volume of sales decreases, the efficiency of the audit method relative to the observation method tends to increase; this is an important factor to remember when making a choice of methods.

Projection to Total Using the Observation Method

As originally construed, one objective of the study was to measure the effectiveness with which the observation procedure could measure week-to-week changes in sales as compared with the standard auditing procedure. But at the middle of the first marketing season, plans were altered to provide a test of the observation technique for estimating *total* sales of apples and pears. The "A" sample was altered to permit estimates of total sales for a week by selecting observation hours during Monday, Tuesday, and Wednesday as well as during the latter half of the week. The "B" sample was maintained on the old basis, namely, observations during Thursday, Friday, and Saturday.

The redesign of the "A" sample made use of 3 by 3 latin square designs. Individual stores were designated as rows, weeks were designated as columns and slow, medium, and busy periods as "treatments." This 3 by 3 arrangement was repeated 10 times over the sample of 30 stores and extended over a period of 12 weeks. In total, this amounted to 40 latin squares (fig. 1). Furthermore, estimates of variation among observation hours within stores and within slow, medium, and busy periods were obtained by taking an additional observation in each of 3 stores each week, as indicated in figure 1. These 3 additional observations were rotated over all 30 stores during the 12 weeks, February through April. Note that the same latin square design was repeated for each 3-week period of the 12 weeks. Thus, on the 1st, 4th, 7th, and 10th weeks, the same sample hours in each of the stores were included in the sample.

FIGURE 1.—*Diagramatic representation of the redesigned observation sample*

Store No.	Week							
	1	2	3	4	5	6	7	Etc. 12
1-----	B/B	S	M	B	S	M		
2-----	S/S	M	B	S	M	B		
3-----	M/M	B	S	M	B	S		
4-----	B	S/S	M	B	S	M		
5-----	S	M/M	B	S	M	B		
6-----	M	B/B	S	M	B	S		
7-----	M	B	S/S	M	B	S		
8-----	B	S	M/M	B	S	M		
9-----	S	M	B/B	S	M	B		
				B/B	S	M		
				S/S	M	B		
				M/M	B	S		
30----								

B=Busy customer activity period.
M=Medium customer activity period.
S=Slow customer activity period.
B/B=Two observations in the same store.

For each week, results from two types of estimators were compared:

(1) An "unbiased" projection, to totals for the 30 stores, of apple and pear sales during the observed hours using the reciprocals of the

sampling rates for the slow, medium, and busy customer activity strata; namely,

$$X'=\sum_{i=1}^{R} W_i \sum_{j=1}^{n_i} X_{ij}$$

in which X'=estimated total sales for the 30 stores,
X_{ij}=sales during jth observation period in the ith customer activity stratum,
W_i=reciprocal of the sampling rate in the ith stratum,
n_i=number of observation hours in ith stratum,
and R=the number of strata.

(2) A ratio estimate, namely, apples sold divided by the customer count during the corresponding hours, multiplied by total number of customers:

$$X''=Y\frac{\sum_i W_i \sum_j X_{ij}}{\sum_i W_i \sum_j Y_{ij}}$$

in which Y_{ij}=customer count observed for the jth observation period in the ith stratum,
and Y=total number of customers. (See text below.)

During the preliminary survey, data on number of customers per week were obtained for each sample store. Y is the total of these numbers of customers for the 30 sample stores. Some "statistical inconsistency" in the ratio estimate was recognized as the sum of Y_{ij} over the population of observation hours in the 30 stores was not necessarily equal to Y, because of time differences and because Y was not a count of customers based on the same procedure that was used to obtain the counts Y_{ij}. However, as current values of ΣY_{ij} summed over the population of observation hours in the 30 stores were not available, the ratio estimator as defined above was the only ratio method of estimation that could be used.

In tables 2 and 3, the "unbiased" and ratio estimates are compared with corresponding estimates from the audit method. Although the ratio estimates, as compared with the unbiased estimates, had the lowest sampling error for estimating total sales and their average deviation from the audit totals was less, the evidence in favor of the ratio estimate as defined above is not conclusive for reasons indicated in the previous paragraph. The results suggest, however, that possibilities of developing a "consistent" ratio estimator should be considered. For example, with information on "checkouts" for the observation hours and corresponding information on "checkouts" by the sample stores for the week, a "consistent" ratio estimator could be developed.

TABLE 2.—*Average percentage deviation of the projected observation estimates from the audit total sales*

Method of estimating	Commodity	
	Apples	Pears
	Percent	*Percent*
Unbiased estimate	33. 3	25. 0
Ratio estimate	12. 6	3. 5

TABLE 3.—*Average coefficients of variation of the projected observation estimates*

Method of estimating	Commodity	
	Apples	Pears
	Percent	*Percent*
Unbiased estimate	18. 1	21. 3
Ratio estimate	12. 5	16. 1

Efficiency of Stratification

The efficiency of stratification as discussed below, refers to the observation method and a single stage sampling plan, the observation periods being the sampling units. That is, the concept is that of a population of sampling units (all possible observation periods) which may be stratified in one of two ways: (1) By slow, medium, or busy sales periods, disregarding the store that any particular sampling unit may belong to, or (2) by size of store, which means that all observation periods within stores of a given size group would constitute a stratum.

Such information is not available for all stores prior to sampling, but it can be obtained by making a preliminary survey using a large sample of the kind used for this study. The analysis assumes that the size of the sample for the prelimi-

Table 4.—*Gains resulting from stratification, proportionate allocation*

Commodity	Stratification by	
	Volume	Customer activity
	Percent	*Percent*
Apples	23	33
Pears	16	30

Table 5.—*Gains resulting from stratification, optimum allocation*

Commodity	Stratification by	
	Volume	Customer activity
	Percent	*Percent*
Apples	40	58
Pears	24	64

nary survey relative to the size of the operating sample is large. In fact, the stratum sizes were assumed to be known without error. To the extent that this is not true, the efficiency of stratification as stated below is overestimated.

Gains due to stratification were computed for each of 12 weeks for the quantity of apples and pears sold, using the following formula,

$$\frac{\sigma_w^2+\sigma_b^2}{\sigma_w^2}-1$$

in which σ_w^2=within stratum variance,
and σ_b^2=between stratum component of variance

The average gains from stratification are shown in table 4. They are based on proportionate allocation of the sample and pertain to estimates of amounts sold and not week-to-week changes. The effect of stratification would be less for week-to-week changes.

The stratification of observation periods by size of store produced a 23-percent gain in efficiency for apple sales compared with a 16-percent gain for pears. There were three size-of-store strata based on dollar volume:

Under $6,000 per week.
$6,000–$20,000 per week.
Over $20,000 per week.

The gains in efficiency from stratification of the observation periods by slow, medium, and busy customer activity were more pronounced than for the size-of-store criteria. The improvement in sampling efficiency was 33 percent for apples and 30 percent for pears. As in the case of store size, the information used for stratification by customer activity was collected from store managers as part of the preliminary survey.

Gains from optimum allocation of the sample to strata are shown in table 5 for apples and pears, for store size, and for customer activity stratification. Optimum, as compared with proportionate allocation, provides a substantial improvement in efficiency. The gains for optimum allocation by customer activity are approximately 60 percent for both apples and pears compared with gains of about 40 percent for apples and 24 percent for pears when using dollar volume.

A system of stratifying observation periods involving a combination of store size and customer activity should be better than either of the methods described. This would be expected because the rating of hours by customer activity is a "within-store" comparison that does not reflect differences in size of store.

Effect of Length of Observation Time on Sampling Efficiency

As the statistical efficiency of a sample is related to the size of sampling unit, in this case length of observation period, the observation data were recorded separately for the first half-hour and the second half-hour of each observation period. This permitted the estimation of variances of sales for apples and pears by slow, medium, and fast customer activity periods, and for half-hour and 1-hour periods. These estimates, which are summarized in table 6, were derived from the replicated observations in the modified "A" sample. The variances for half-hour and 1-hour periods were then plotted against the length of observation period on a log-log chart and the points joined to provide estimates of the variance for other lengths of observation. The assumption underlying this

Table 6.—*Within-store variances of sales for half-hour and hour periods, by customer activity*

Commodity and customer activity	Length of observation period	
	One hour	One-half hour
Apples:		
Slow	11. 8	8. 1
Medium	32. 0	22. 0
Busy	152. 5	90. 2
Pears:		
Slow	2. 0	1. 7
Medium	3. 5	2. 6
Busy	13. 8	8. 3

Table 7.—*Cost by Length of observation*

Length of observation period	Cost of travel and observation
Hours	*Dollars*
0. 50	3. 80
. 75	4. 15
1. 00	4. 50
1. 50	5. 20
2. 00	5. 90
2. 50	6. 60
3. 00	7. 30

procedure is that the log of the variance is linearly related to the log of the length of observation[1] (size of sampling unit).

To determine the optimum length of observation per dollar, data on certain components of cost are needed also. Within the range of the sample sizes that we anticipated might be used, it appeared that the variable part of the field costs would be adequately approximated as a linear function of the cost of observation and of travel between stores which are described in the next section of this report.

Using the variance and cost functions, optimum sizes of clusters were developed for each of the three customer activity periods. The procedure used was to assume a fixed total dollar cost that was to be expended on the survey operation. Based on the fixed cost of travel and observation, the number of stores that could be visited was then computed for half-hour, hour, hour and a half, 2-hour intervals, and so on. The schedule of costs used is shown in table 7.

From the derived number of stores and the variance read from the plotted lines, the relative sampling error was then computed. The lowest sampling error derived in this way determined the optimum length of observation. Table 8 shows the estimated optimum length of observation for the slow, medium, and busy customer activity periods for apples and pears.

Since this analysis is based on two points for a variance function, the estimates of optimum length are undoubtedly subject to considerable error, especially the estimates that involve a high degree of extrapolation. But as the neighborhood of optimum usually ranges over a wide area, the derived optimum sizes are probably not far from the true optimum.

Cost Analysis

A cost analysis was prepared for the audit and observation operations, exclusive of overhead office costs and supervisory costs, as these are constant elements that do not vary materially with changes in sample size.

It may be noted in table 9 that the average travel cost per observation, $1.84, was 21 cents higher than the average cost of travel for an audit visit. This is reasonable when it is recognized that visits for observation purposes are prescribed in advance. In the case of the audit operation, the auditor can make his rounds in the most convenient and expeditious fashion. The miscellaneous costs per store visit, which averaged 51 cents, represent costs of recopying, correcting, and preparation of forms for mailing, and are approximately the same for both audits and observations. The average cost per observation for 1-hour periods was $1.62. This of course is constant over all sizes of stores since the interviewer must observe for 1 hour irrespective of the volume of traffic in the particular store.

The direct audit costs were analyzed in terms of total dollar sales volume of the store. The smaller stores with an average weekly volume of less than $6,000 cost on the average 95 cents, compared with $1.23 for a larger store with a weekly dollar volume of $20,000 or more. These average cost figures include the cost of interviewer time, less cost of travel and miscellaneous expenses. The average number of minutes for travel, observation, and audit are shown in the adjoining column.

---

[1] A number of investigations have found this assumption to be satisfactory. Jessen, R. J., *Statistical Survey for Obtaining Farm Facts*, Iowa Agricultural Experiment Station, Res. Bul. 304 (1942). Mahalanobis, P. C., on Large Scale Sample Surveys, Phil. Trans. Roy. Soc. London (1944).

Test of Design Findings

For the second marketing season or phase of this study, the sample was redesigned in view of the experience and information gained during the first season. As before, a preliminary survey was made. The method consisted of:

(1) Selecting an area sample of the metropolitan area using Sanborn clusters and Business Enumeration Districts.[2]
(2) Enumerating of food stores in these clusters.
(3) Making a comparison of these enumerated food stores with food stores on route lists published by a newspaper.
(4) Eliminating from the area sample those stores that were on the route lists, and
(5) Selecting a sample of food stores from the route lists and from stores remaining in the area sample after step (4).

In the sample for the preliminary survey there were 612 food stores, including those selected from the route list and the ones in the area sample that were not on the route list. Information collected from these stores included:

(1) Whether or not the outlet carried fresh products.
(2) Total dollar volume of sales.
(3) If self service, the number of check out counters.
(4) Opening and closing times and a classification of all hours as slow, medium, or busy periods.
(5) Estimated average weekly customer count.

On the basis of these sample data, and the variances from the previous year's study, the audit and observation samples were redesigned.

Table 8.—*Optimum length of observation time*

Customer activity	Commodity	
	Apples	Pears
	Hours	*Hours*
Busy	0.75	0.75
Medium	1.50	3.00
Slow	2.50	5.50

Table 9.—*Field costs*

	Dollars	Minutes
Average travel cost per observation	1.84	82
Average travel cost per audit visit	1.63	73
Average miscellaneous cost per store visit	.51	23
Average in-store cost per observation	1.62	72
Average cost per store audit, by weekly dollar volume:		
Under $6,000	.95	42
$6,000–$10,000	1.05	47
$10,000–$20,000	1.06	47
$20,000 or more	1.23	55

As there was less emphasis on a direct comparison of the audit and observation methods during the second marketing season, both methods were not used in the same stores.

Considering costs and sampling error, the first year's experience suggested use of audits in small stores and observations in large stores. This was because, for large stores as compared with small stores, the cost of an audit relative to an observation was higher and the statistical efficiency of audits relative to observations was less.

Self-service did not occur in stores that had less than $6,000 in total weekly sales. This appeared to be a good boundary for determining the subuniverses of small stores and of supermarkets for which the audit and observation procedures would be used. For the operating sample, 100 stores were selected from the 612 enumerated in the preliminary survey. Fifty-nine of the sample stores were audited and observation procedures were used in 41.

---

[2] A Sanborn cluster is a group of retail stores identified on Sanborn maps published by the Sanborn Map Co. The Business Enumeration District was a unit of work developed by the U.S. Bureau of the Census for its 1948 Census of Business. Enumeration districts were used where Sanborn map coverage was not available.

Table 10.—*Sample allocation of observation periods*

	Customer activity strata		
	Busy	Medium	Slow
Total number of hours (all 41 stores)	924	608	658
Length of observation period in hours	.75	1.5	2.5
Number of sampling units, N_i	1,232	405	263
σ_i	9.06	3.67	1.94
$N_i\sigma_i$	11,162	1,486	510
Theoretical allocation of sample	34.8	4.6	1.6
Actual allocation of observation periods			
Size of store (dollars per week):			
6,000–10,000	10	1	1
10,000–20,000	7	3	1
20,000–30,000	2	1	1
30,000–40,000	4	1	-------
40,000 or more	6	1	2
Total	29	7	5

After deleting stores in the preliminary survey that did not have fresh produce counters and stratifying by total weekly sales, the sample was allocated to size groups approximately in proportion to the product of the square root of average dollar sales per store in a stratum and the number of stores in that stratum derived from the preliminary survey.

With reference to the selection of observation periods in the 41 observation stores, the previous year's study indicated that appreciable gains in efficiency could be achieved by (1) stratifying the time periods by slow, medium, or busy customer activity, and (2) varying the sampling rates and the length of observation period by the three customer activity strata.

As described in the first part of this article, it was determined that the optimum length of observation period was about three-fourths of an hour for busy periods, 1½ hours for medium periods, and 2½ hours for slow periods.

To approximate an optimum allocation of the observation periods in the 41 stores, the hours reported as busy, medium, or slow were converted to numbers of sampling units by dividing the reported hours by the length of observation period given above. Using average variances derived from the previous year's study, an allocation of 41 observation periods was then determined as shown in table 10. The actual allocation by customer activity strata differs from theoretical because (1) it was thought that the estimates of variance for medium and slow periods might be subject to more error than the estimate for busy periods, as they were further removed from the actual observed cluster size, and (2) the additional observations in these two strata would provide a basis for verifying the estimates of variance.

Since the measurement of week-to-week change in the volume of apple sales was of considerable importance, the sample of audit stores and the observation periods selected from the 41 observation stores were held fixed throughout the second marketing season. Substantial gains in efficiency in measuring change can be achieved by holding the sample fixed from week to week.

Collection Procedure

Audits were conducted in the 59 audit stores on Monday through Wednesday of each week throughout the duration of the study using the standard procedure of recording opening and closing inventories, spoilage, returns, and purchases over the week just past. The weekly sales for each store were then derived from the equation: Opening inventory, plus purchases minus spoilage, minus returns, minus closing inventory, equals sales.

The observation sample was distributed throughout the week from Monday to Saturday. The observer recorded the purchases of apples by varieties and simultaneously recorded the number of customers passing the fresh produce counter during the observation period.

Estimates of Total Apple Sales

An estimate of the total sales of apples in Philadelphia was made each week by preparing separate estimates from the audit and observation samples and adding the results together. For the audit sample, the mathematical formulation of the estimate was as follows:

$$X' = \sum_{i=1}^{R} \frac{M_i}{m_i} \sum_{j=1}^{m_i} X_{ij}$$

in which R=number of audit strata,
M_i=number of stores in the ith stratum, which is estimated from the preliminary survey,
m_i=number of sample stores in the ith stratum,
X_{ij}=apple sales in the jth store of the ith stratum.

In the case of the observation data, an aggregate ratio type of estimate was used:

$$X'' = Y \frac{\sum_{i=1}^{S} \frac{N_i}{n_i} \sum_{j=1}^{n_i} X_{ij}}{\sum_{i=1}^{S} \frac{N_i}{n_i} \sum_{j=1}^{n_i} Y_{ij}}$$

in which Y=estimated total number of customers in all stores in the Metropolitan Market in the subuniverse that was sampled by the observation method,
S=number of observation strata,
N_i=total number of sampling units (possible observation periods) in the ith stratum for the 41 stores (see table 10),
n_i=sample number of sampling units (observation periods),
X_{ij}=observed apple sales for the jth observation in the ith stratum,
Y_{ij}=observed customer count for the jth observation period in the ith stratum.

This formula provides a weighted estimate of apple sales based on the reciprocal of the sampling rates for the sample observation periods, divided by a similarly weighted customer count for the same periods. This ratio was then multiplied by the estimated total number of customers in the observation subuniverse, which was developed from the preliminary survey, to obtain an estimate of total apple sales for the observation subuniverse. The observation and audit estimates combined provided an estimate of the total apple sales through retail stores.

Reliability of Estimates of Totals and Change

Standard errors were computed for estimates of total retail apple sales for four weekly periods. These estimated errors ranged from 3.8 to 7.7 percent of the estimated total sales, the average being 6.3 percent. To estimate the standard error of week-to-week ratios of sales, 7 selected pairs of weeks were used. The estimates ranged from 0.061 to 0.094, the average being 0.074.

Some simplifications or approximations were made in the establishment of a procedure for estimating the standard errors quoted in the previous paragraph. One of these approximations was the collapsing of strata. For example, though eight dollar-volume strata were used in selecting the audit stores, only three were used for computation purposes. Similarly, in the case of the observation sample, three strata were used in selection—busy, medium and slow, but only two—busy and a combined medium and slow strata were used for computational purposes. These simplifications had the effect of overstating the sampling error.

A second approximation was the omission of the error in the expansion factors. The design as outlined above is a double sample. The expansion factors used in the projection formula were determined from the preliminary study and are therefore subject to sampling error. But the error contribution of the strata weights are believed to be small and may compensate for the collapsing of strata in the error computation.

The variance function used in computing these sampling errors was as follows:

$$\text{Var. of } (X'+X'') = \sum_{i=1}^{R} \left(M_i^2 \frac{M_i - m_i}{M_i - 1} \frac{V_i}{m_i} \right) + Y^2 \sum_{i=1}^{S} W_i^2 \frac{1}{n_i} \frac{\overline{X}^2}{\overline{Y}^2} \left(\frac{V_{xi}}{\overline{X}^2} + \frac{V_{yi}}{\overline{Y}^2} - \frac{2 \operatorname{cov} X_i Y_i}{\overline{X}\,\overline{Y}} \right)$$

in which
R=number of audit strata,
S=number of observation strata,
V_i=variance of audit sales in ith audit stratum,
M_i=number of stores in the ith audit stratum,

m_i=number of sample stores in the ith audit stratum,

n_i=sample number of stores or observation periods in the ith stratum,

Y=estimated total number of customers in the observation subuniverse,

$W_i=\frac{N_i}{\Sigma N_i}$=The relative weight of each observation stratum to the total observation subuniverse,

$\overline{X}$=average observed sales per observation period weighted for the different sampling rates in each observation stratum,

$\overline{Y}$=weighted average customer count corresponding to $\overline{X}$,

V_{xi}=variance of observed sales in ith observation stratum,

V_{yi}=variance of customer counts in ith observation stratum,

cov X_i Y_i=covariance of observation sales and customer counts for ith observation stratum.

Summary

One of the original purposes of this study was to compare the efficiency of the audit and observation techniques. But as either technique may provide some information that cannot be obtained by the other, the choice is more than a matter of comparative statistical efficiency and cost. When the average volume of sales per observation period is low, either because a store is small or because the commodity is a low-volume item, the sampling error for the observation method will be comparatively large. Considering statistical efficiency and cost, the observation method is most feasible in large stores. A combination of the two methods, audit in small stores and observation in large stores, was tried successfully during the second year this project was in effect.

Some variations of the audit and observation methods are discussed briefly in the following article.

Observation Techniques in Store Auditing

By Solomon Dutka and Lester R. Frankel

This paper describes some variations of audit and observation methods that are discussed in the preceding article, "Observation and Audit Techniques for Measuring Retail Sales."

ONE OF THE MAJOR TOOLS used in deriving marketing data at the retail level is the standard retail store audit procedure. Classically, this procedure consists of (1) selecting a representative sample of outlets for auditing, (2) securing their cooperation, (3) taking a physical inventory of the item under study and (4) collecting the purchase invoices maintained by the retailers. The retail sales for the items are then derived by the formula: Opening inventory plus retailer purchases minus closing inventory equals sales for the period.

For a variety of reasons, other procedures for obtaining similar types of information (as well as data not obtainable using the standard procedure) which can be broadly classified as "observation" techniques have been utilized. They include all procedures used in obtaining data on either sales, purchases, or inventories other than through the use of the conventional physical inventory and purchase invoice procedure.

The major reasons for the newer techniques are the limitations of the physical audit mechanism itself, including problems of response error and cost considerations. For example, in the accompanying article, "Observation and Audit Techniques for Measuring Retail Sales," reference is made to the use of the "observation" technique in the measurement of retail sales of apples and pears. In this study, conducted for the U.S. Department of Agriculture, a comparison was made of estimates of retail sales of fresh produce (apples and pears) using the standard audit procedure with similar estimates using the observation technique. In this latter procedure, retail purchases by customers were observed and recorded for time-segment samples within stores.

In comparing the results of the two procedures directly, some of the limitations of the physical audit mechanism were apparent. For one, the physical audit and the observation audit *cannot* measure sales data for the same time period because, in conducting the standard audit, all stores cannot physically be audited at the same time and on the same day at uniform time intervals. Even if a large enough field staff can be mustered, retailers generally prefer auditors to conduct their activities so as to cause a minimum of disturbance to the store's general activities. During peak periods, such as Christmas and Easter, this problem becomes even more acute.

Another factor in these comparisons is that although the observation technique used in this procedure recorded actual sales to consumers, the physical audit measures the movement of goods off the premises, which is assumed to be equal to retailer sales. Thus, the physical audit includes spoilage and returns and transshipment (unless specifically recorded) as retail sales.

For other consumer items, the problem of pilferage is especially acute in certain types of retail outlets amounting to a significant item in their retail profit picture. Here again, the physical audit procedure mechanism is subject to an overstatement of retail sales as compared with the observation procedures. Also, certain types of products, such as pharmaceuticals, do not allow for an accurate physical count of inventory.

Failure to achieve full cooperation in conducting the standard audit procedure can cause significant biases in estimating sales data. Failure on the part of the retailer to retain all invoices during the audit interval or to keep records of returns and transshipments can likewise cause biases. Also, in certain types of establishments, the bulk of the retailer purchasing is done on a cash basis with either no invoice recording or simply a summary bill with no description. Finally, it can be shown that for certain high turnover items where substantial time is involved in actually obtaining complete counts of physical inventories and invoices (such as with soft drinks and frozen foods), an optimum solution taking into account the cost and variance functions might suggest the observation technique.

Some additional types of "observation" techniques that have been used follow:

Inventory depletion.—In an effort to derive retail sales data from among stores that cannot provide full cooperation for the standard audit procedure, a system is utilized wherein counts are made of the retail counter or shelf inventory during a cross section of time periods when the store is open. In effect, the procedure consists of recording the depletion in retail inventory at the counter or shelf between restocking periods as a measure of purchases of these items by consumers. Within the cost limitations, this procedure can be made more efficient statistically by oversampling certain of the hours or days during which the store is busiest. This procedure has been found to be particularly useful for consumer items with fairly high turnover. One qualification of this technique is that, though estimates of retail sales can be obtained, it permits no direct method for obtaining data on retailer purchases from wholesalers or on retailer inventories, which might be important for certain products.

Direct observation of sales.—Another variation of the observation technique is that used in measuring retail sales of apples and pears, as previously referred to. This procedure has particular merit for application to measuring the retail movement of fresh produce wherein it might be physically too difficult to utilize the "inventory depletion technique" or to obtain inventory information using the standard audit procedure.

The purchase invoice technique.—Another variation of the observation technique used to derive estimates of retail sales of certain products is the "purchase invoice technique." It is particularly relevant to items that have a relatively large turnover compared with their inventory at any time. Retailer purchase invoices as estimates of retailer sales can likewise be used for other high-turnover products, such as fresh milk and ice cream. There is a constant flow of deliveries of these products and thus little need for any substantial retail inventory.

Direct measures of sales.—Frequently the problem of partial cooperation has been encountered among retailers in auditing products that have high unit cost. Often retailers will permit the collection of certain types of information but not others. In auditing jewelry and cameras, a variation of the warranty card procedure has been employed with the card being filled out at the retail counter. As each consumer purchase is made, the retailer fills out a form attached to the sales slip, then this form is mailed directly to the surveying company. In effect, this provides a direct measure of the sales of these items without having to take complete inventory and collect purchase invoice data.

In providing clients with estimates of retailer sales of gasoline, the procedure of measuring gallons pumped during the audit interval has been used as a measure of retail sales. This does not require an estimate of the inventory at any time within the pumps themselves.

For certain products, such as ethical drug products, costs of collecting inventory information are prohibitive. In order to measure retail sales of prescription items, a procedure of sampling time segments from the pharmacist's prescription book has been used. By sampling within the audit interval and obtaining a copy of the prescriptions recorded, total estimates of sales of prescription items can be made.

In summary, collection of data on retail sales, purchases, or inventories at the retail level through procedures other than the standard physical audit procedure, one which we have labeled "observation" techniques, came into being because of the specific purposes of the audit, the product class or type to be audited, and cost considerations.